



# Articulated Shape Matching by Robust Alignment of Embedded Representations

Diana Mateus, Fabio Cuzzolin, Radu Horaud, Edmond Boyer

## ► To cite this version:

Diana Mateus, Fabio Cuzzolin, Radu Horaud, Edmond Boyer. Articulated Shape Matching by Robust Alignment of Embedded Representations. 3DRR 2007 - IEEE Workshop on 3D Representation for Recognition, Oct 2007, Rio de Janeiro, Brazil. pp.1-8, 10.1109/ICCV.2007.4408833 . inria-00590238

**HAL Id: inria-00590238**

**<https://inria.hal.science/inria-00590238>**

Submitted on 3 May 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Articulated Shape Matching by Robust Alignment of Embedded Representations

Diana Mateus   Fabio Cuzzolin   Radu Horaud   Edmond Boyer  
PERCEPTION group, INRIA Rhône-Alpes, France  
{mateus, cuzzolin, horaud, boyer@inrialpes.fr }

## Abstract

*In this paper we propose a general framework to solve the articulated shape matching problem, formulated as finding point-to-point correspondences between two shapes represented by 2-D or 3-D point clouds. The original point-sets are embedded in a spectral representation and the actual matching is carried out in the embedded space. We analyze the advantages of this choice as well as the reasons for which the task remains a difficult one. In particular, we show that although embedded-space matching still has intrinsic combinatorial difficulties, it can be solved by searching for an optimal orthogonal transformation that aligns the two shape embeddings. Relying on the model based clustering formalism, we propose a probabilistic formulation which casts the matching into an EM algorithm. Outliers are properly handled by the algorithm and a simple strategy is adopted to initialize it. Experiments are performed with three embedding methods (Isomap, LLE, and Laplacian embedding) and with 3-D voxelsets representing a human-motion sequence.*

## 1. Introduction

Shape matching for establishing shape similarities is a central issue in computer vision. It can be used to find shape classes for object recognition, to track objects over time, to build spatio-temporal representations useful for shape modelling, for action and/or gesture recognition, etc. Furthermore, shape matching has recently been found to be useful in other domains such as biochemistry allowing one protein to be aligned with another [22], and in astronautics where the attitude of rockets or satellites can be estimated by matching the shape of a stellar constellation acquired with an onboard camera with an empirical star catalog [17].

Although there are methods available both for rigid objects and for objects with deformable surfaces, articulated shape matching remains a difficult problem. There is no simple relationship between the underlying kinematic pa-

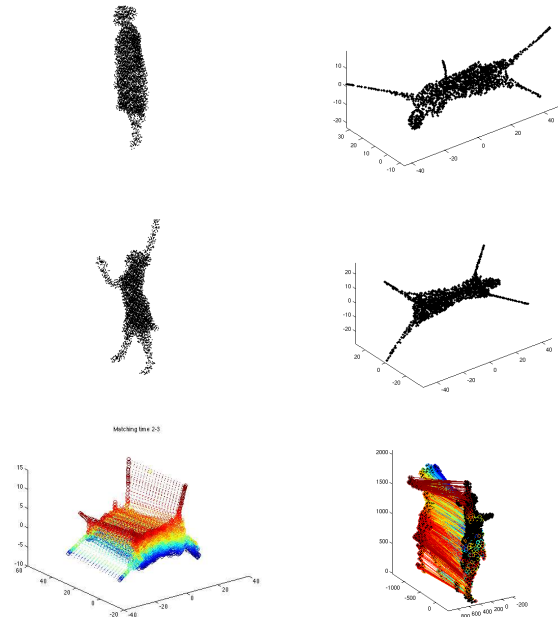


Figure 1. Illustration of our method. Each shape to be matched is described by a set of 3-D points (left of top and middle rows). Spectral embeddings of the sets (right of top and middle rows). The eigenspaces associated with the two embedded representations are robustly aligned which results in a one-to-one correspondence between the two sets of points (bottom row).

rameterization and the information that is actually observable in the 2-D or 3-D visual data. Rigidity (and hence isometry) is only locally preserved and knowledge about how an articulated shape is split into rigid pieces is not necessarily available.

In this paper we address the following problem: Let an articulated shape be represented by a set of points (2-D silhouettes, 3-D visual hulls, meshes, voxels, etc.) We consider two instances of this set of points in two different poses (for example, at two different times) and we want to establish a one-to-one assignment between the two point sets, e.g., Figure 1. We consider difficult situations when

the data are corrupted by occlusions (points present in one set may be absent in the other set), missing data, outliers, and noise.

A recurrent idea for non-rigid shape description and matching is the use of dissimilarity matrices with entries corresponding to pairwise distances between points of a single shape. In a very simplified manner, matching two shapes represented this way implies minimizing the discrepancy between the two dissimilarity matrices over all possible pairings. [21], [5, 6], [11], [15].

In order to describe articulated shapes, it is desirable to define the pairwise distance measure in such a way that local isometric properties are preserved while allowing the matrix to be “globally invariant” to shape deformations. The dissimilarity matrix can be interpreted as the adjacency matrix of a weighted graph: nodes describe points and weighted edges describe links to neighboring points. In this context, the problem of shape matching can be cast as graph isomorphism, namely finding a node-to-node assignment that minimizes the Frobenius norm of the difference between the two dissimilarity matrices. A modern approach to graph isomorphism is to define probabilistic assignments and to use variations of the EM algorithm [12, 9, 16]. These deterministic algorithms guarantee that the error decreases at each iteration but cannot escape from local minima, and so a proper initialization is crucial. Furthermore, these graph-matching methods usually fail to match very large sets of points, to deal with graphs with very different sizes, and to compare very different graphs, i.e., graphs that contain many spurious nodes.

Spectral graph theory [7] unfolds several interesting properties of Shape Matrices: its eigenvalues and eigenvectors characterize the spectrum and the modes of the graph. The eigenvalues alone may be used to compare shapes, although distinct shapes may lead to the same eigenvalues [14, 18] More interesting, the eigendecomposition itself has been used to build a closed-form solution for matching weighted graphs of equal size [24], and to solve for image matching problems [20], [21], [5, 6].

It is generally conjectured that the embedded representation obtained through the spectral analysis of the dissimilarity matrix is appropriate for shape description and analysis. An important reason for this is the fact that the matrix encodes pairwise geometric invariants, which are also preserved in the embedding space. Techniques used for non-linear manifold learning, have exploited this principle for dimension reduction: building small dimensional representations of manifolds only known through samples that lie in a high dimensional space. The main idea is that by using spectral analysis on adequately designed dissimilarity matrices one can constrain the structure (global or local) of the observed space to be preserved in the embedding

space. Indeed, methods such as Isomap [23], locally linear embedding (LLE) [19], Laplacian embedding [2], etc., use some more elaborate notions of distances and neighborhoods to build a dissimilarity matrix, but converge at some point to the computation of the eigenvalues and their associated eigenvectors, to finally retain a reduced eigenspace spanned by the most significant (the largest or the smallest) eigenvalues/ eigenvectors pairs.

Let  $n \times 2$  or  $n \times 3$  be the dimension of the original cloud of points, and  $n \times n$  be the size of the dissimilarity matrix where  $n$  is the number of points associated with a shape. The former can be embedded into a vector space of dimension  $k$ ,  $k < n$  whose canonical basis corresponds to the eigenvectors associated with the  $k$  largest eigenvalues of the matrix. It is therefore possible to project the set of points onto this space. Two similar shapes should share the same embedded representations due to the fact that their matrices capture only local isometric properties and therefore matching shapes in this space is only weakly affected by articulated deformations. It is therefore tempting to conjecture that shape matching is now reduced to eigenspace alignment [11], [15]. Nevertheless, several problems arise:

- Eigenvectors are only defined up to sign which introduces  $2^k$  possible alignments between the two eigenbases: This will be denoted as the *sign-reversal ambiguity*;
- The ordering of eigenvalues is not reliable. This has several reasons. First, the shape itself may yield several identical elongations and hence several similar eigenvalues. Second, there may be eigenvalues with large algebraic multiplicity. Third, there may be numerical instabilities that do not allow proper ordering. There are  $k!$  such possible “switches” between the eigenvalues: This will be denoted as the *eigenvalue-ordering ambiguity*.
- The two sets of points may never match exactly, owing to spurious points, missing, bad, and/or noisy data, etc. As a consequence, the embedded spaces cannot be perfectly aligned and there is a *residual* transformation (rotation and/or translations) between the two bases: This will be denoted as the *rigid-motion ambiguity*.

To summarize, there are  $2^k k!$  possible alignments between the two eigenbases, and each one of these alignments is defined up to a rotation and translation. Table 1 summarizes the number of possible alignments as a function of  $k$ . This is a truly combinatorial problem because the number of solutions (or configurations) increases dramatically with the dimension of the embedded space, and there is no way

to guide the search through the space of possible configurations. It is interesting to notice that, although the sign-reversal and eigenvalue-ordering ambiguities have been reported in [21], none of the methods mentioned above properly addresses these issues (although some heuristics have been proposed in [15, 4]).

Table 1. The number of possible shape alignments as a function of the dimensionality of the eigenspace being considered.

| $k$        | 2 | 3  | 4   | 5    | 6     |
|------------|---|----|-----|------|-------|
| alignments | 8 | 48 | 384 | 3840 | 46080 |

One may argue that it is sufficient to consider a low-dimensional embedded space and to perform an exhaustive search. Indeed, is practical for  $k \leq 3$  provided that there is a reliable and efficient way to test each solution. However, in typical cases the dimension of the embedded space is higher than that. Although the literature is abundant with experiments aimed towards the guess of the dimensionality of the embedded space and its exploration for clustering, there is little information about the choice of  $k$  when the task is to align shapes. In [15] it is claimed that  $k = 6$  yields satisfactory mesh correspondence and good results are obtained with a greedy method. As already mentioned, in [14] it has been observed that two similar shapes will always yield comparable eigenvalues but the reverse is not always true. We conclude that for reliable spectral shape matching, the dimension of the embedded space should not be too small, most certainly in the range  $5 \leq k \leq 10$ , as reported in [15].

This paper has the following contributions. Articulated shape matching is formulated as the problem of finding a one-to-one correspondence between two point sets embedded in some spectral domain. We show that this could and should be done by searching for an optimal *orthogonal transformation* between their spectral descriptions. This transformation encodes the three spectral ambiguities mentioned above in a unified manner. We introduce a probabilistic formulation allowing to model the point matching process as an EM algorithm that treats the correspondence in a *non symmetric* manner. Since EM is a deterministic minimizer, it requires proper initialization. We discuss possible initialization strategies.

## 2. Problem formulation

We denote by  $\mathbf{X} = \{\mathbf{X}_i\}_{1 \leq i \leq n}$  and by  $\mathbf{Y} = \{\mathbf{Y}_j\}_{1 \leq j \leq m}$  the two sets of points. We seek a one-to-one correspondence between the two point sets. Without loss of generality we consider 3-D points and let  $n \geq m$ . Let  $A$  and  $B$  be the dissimilarity matrices associated with the

point sets  $\mathbf{X}$  and  $\mathbf{Y}$ . Taking as example the *Laplacian embedding* [2], an entry  $A_{ij}$  of such a matrix is computed with  $A_{ij} = \exp(-d^2(\mathbf{X}_i, \mathbf{Y}_j)/2\sigma^2)$ . We also define a diagonal matrix  $D$  such that  $D_{ii} = \sum_j A_{ij}$ . We seek the eigenvalues and generalized eigenvectors that satisfy  $\det(A - \lambda D) = 0$  and  $A\mathbf{U} = \lambda D\mathbf{U}$ , where the vector columns  $\mathbf{u}_i$  of the  $n \times n$  matrix  $\mathbf{U}$  are solutions of the equation  $(A - \lambda_i D)\mathbf{u}_i = \mathbf{0}$ . There are well known methods to extract the  $k$  most significant eigenvectors. We retain  $U_{n \times k}$  whit columns corresponding to the most significant eigenvectors  $\mathbf{u}_1, \dots, \mathbf{u}_k$ , and whose rows are the coordinates of the points in the embedded  $k$ -dimensional space,  $\mathbf{x} = \{\mathbf{x}_i\}_{1 \leq i \leq n}$ . Similarly we denote by  $\mathbf{y} = \{\mathbf{y}_j\}_{1 \leq j \leq m}$  the embedded coordinates of the second set of points obtained from the dissimilarity matrix  $B$ .

Other than the *Laplacian embedding*, one can obtain similar spectral representations using embedding techniques such as Isomap [23] and local linearly embedding (LLE) [19]. The method that we describe below can be applied to anyone of these embedded representations.

Let  $Q$  be the  $k \times k$  matrix that allows, in principle, the alignment of the two embedded spaces associated with the two sets of points. This matrix is formed by the following multiplication:

$$Q = RPS \quad (1)$$

where  $R$  is a  $k \times k$  rotation (i.e., an orthogonal matrix with  $\det(R) = 1$ ),  $P$  is a  $k \times k$  permutation matrix, and  $S$  is a  $k \times k$  diagonal matrix with entries  $s_{ii} = \pm 1$ . These three matrices encode (from right to left) the sign-reversal, the eigenvalue-order, and the rigid-motion ambiguities. There are  $k!$  possible permutations  $P$  and  $2^k$  possible sign flips  $S$ . Notice that  $P$  is orthogonal, and therefore  $Q$  is an orthogonal matrix as well. The latter can be either a rotation or a reflection. Notice that the translational component of the rigid-motion ambiguity has been left out, this can be easily computed by aligning the centers of mass of the two embeddings.

Points in the two sets can now be aligned according to the formula:

$$\mathbf{x}_{z(j)} = Q\mathbf{y}_j \quad (2)$$

where the correspondence function  $z$  allows to assign a point from the first set to a point from the second set. Since we do not have such an assignment, the embedded alignment can be cast into the following minimization problem:

$$\min_{Q, z} \sum_j \|\mathbf{x}_{z(j)} - Q\mathbf{y}_j\|^2 \quad (3)$$

This formula is similar in spirit with the alignment of two rigid sets of points [1], [27], [13], and with articulated tracking in kinematic space [10]. Therefore, articulated shape

matching has been reformulated as the problem of finding a one-to-one correspondence between two point sets that minimizes an orthogonal transformation between their embedded descriptions. In the next two sections we introduce a probabilistic formulation and its associated algorithm that minimizes a *revised version* of eq. (3).

### 3. Robust spectral alignment

The main difference between our approach and similar ones [8, 16] is that the two sets of points are not treated symmetrically [26], [10]. Indeed, within a probabilistic framework, we define  $\mathbf{x}$  as a set of observed values of an equal number of random variables that will be denoted by  $\mathcal{X}$ . To each random variable  $\mathcal{X}_i$  we associate another random variable  $z_i$  which describes the correspondence. Specifically,  $z_i = \mathbf{y}_j$  means that the observation  $\mathbf{x}_i$  is in correspondence with the point  $\mathbf{y}_j$ . Hence, we refer to the set  $\mathbf{y}$  as the set of model points. Moreover, the notation  $z_i = \emptyset$  means that observation  $\mathbf{x}_i$  is in correspondence with an outlier. Let  $v$  be the volume within which  $\mathbf{y}_j$  is to be expected. The prior probability that an observation point  $\mathbf{x}_i$  is assigned to a model point  $\mathbf{y}_j$  writes  $P(z_i = \mathbf{y}_j) = v/V$  where  $V$  is the volume of the whole space. In order to satisfy  $\sum_j P(z_i = \mathbf{y}_j) + P(z_i = \emptyset) = 1$ , we get  $P(z_i = \emptyset) = (V - mv)/V$ .

We choose to model with different distributions the inliers and the outliers, [26]. The probability of an observation  $\mathbf{x}_i$  to lie in the proximity of  $Q\mathbf{y}_j$ , given that  $\mathbf{x}_i$  and  $\mathbf{y}_j$  are in correspondence, will be described by a Gaussian distribution with covariance matrix  $\sigma I_k$  ( $d$  denotes the Euclidean distance):

$$P_Q(\mathbf{x}_i|z_i = \mathbf{y}_j) = \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-\frac{d^2(\mathbf{x}_i, Q\mathbf{y}_j)}{2\sigma^2}} \quad (4)$$

The probability of an observation given that it corresponds to an outlier will be described by a uniform distribution over the volume of the working space:  $P(\mathbf{x}_i|z_i = \emptyset) = 1/V$ . We are interested in the posterior probabilities, namely the probability that  $\mathbf{x}_i$  corresponds to  $\mathbf{y}_j$  given the observation  $\mathbf{x}_i$ . We will denote these posteriors by  $\alpha_{ij}$ :

$$\alpha_{ij} = P_Q(z_i = \mathbf{y}_j|\mathbf{x}_i) \quad (5)$$

We combine the above expressions for the priors and for the likelihoods in the Bayes formula. We observe that the set  $\{z_i = \mathbf{y}_1, \dots, z_i = \mathbf{y}_m, z_i = \emptyset\}$  is a partition of the event space. Therefore we have:

$$\begin{aligned} P_Q(\mathbf{x}_i) &= \sum_{j=1}^m P(\mathbf{x}_i|z_i = \mathbf{y}_j)P(z_i = \mathbf{y}_j) \\ &+ P(\mathbf{x}_i|z_i = \emptyset)P(z_i = \emptyset) \end{aligned} \quad (6)$$

Finally we obtain the following expression:

$$\alpha_{ij} = \frac{e^{-\frac{d^2(\mathbf{x}_i, Q\mathbf{y}_j)}{2\sigma^2}}}{\sum_j e^{-\frac{d^2(\mathbf{x}_i, Q\mathbf{y}_j)}{2\sigma^2}} + \emptyset_C} \quad (7)$$

With  $\emptyset_C$  representing the *outlier constant*. In order to estimate the transformation  $Q$  we maximize the expectation  $E$  of the logarithm of the joint probability of the observations and their correspondences:

$$F(Q, Q^c) = E_{Q^c}[\log(P_Q(\mathbf{x}, z))|\mathbf{x}] \quad (8)$$

where  $Q^c$  denotes the current estimate of  $Q$ . By supposing that the variables  $\mathbf{x}_i$  and  $z_i$  are independent for all  $1 \leq i \leq n$ , we get:

$$\begin{aligned} P_Q(\mathbf{x}, z) &= \prod_{i=1}^n \left( \prod_{j=1}^m (P(\mathbf{x}_i|z_i = \mathbf{y}_j)P(z_i = \mathbf{y}_j))^{\delta_{y_j}(z_i)} \right. \\ &\quad \left. (P(\mathbf{x}_i|z_i = \emptyset)P(z_i = \emptyset))^{\delta_{\emptyset}(z_i)} \right) \end{aligned} \quad (9)$$

where  $\delta_{y_j}(z_i)$  (respectively  $\delta_{\emptyset}(z_i)$ ) equals 1 when  $z_i = \mathbf{y}_j$  (respectively  $z_i = \emptyset$ ), and 0 otherwise. After some straightforward algebraic manipulations and by grouping constant terms, we obtain:

$$F(Q, Q^c) = -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^m \alpha_{ij} \|\mathbf{x}_i - Q\mathbf{y}_j\|^2 + C^{st} \quad (10)$$

We can further simplify this expression as follows. Let  $\mathbf{w}_j$  be the *average* observation that corresponds to  $\mathbf{y}_j$ , i.e.,  $\mathbf{w}_j = \sum_i \alpha_{ij} \mathbf{x}_i / \sum_i \alpha_{ij}$ . With the notation  $\beta_j = \sum_i \alpha_{ij}$  and by substitution in the above formula, we obtain that the maximization of eq. (8) is equivalent to the minimization of the following criterion:

$$\min_Q \sum_{j=1}^m \beta_j \|\mathbf{w}_j - Q\mathbf{y}_j\|^2 \quad (11)$$

### 4. Algorithm

The formula that we just established can be, in principle, directly used within a non-linear optimization framework. Indeed, the probabilities  $\alpha_{ij}$  (and the weights  $\beta_j$ ) are parameterized by the orthogonal matrix  $Q$  and by the variance  $\sigma$ , and one may well minimize eq. (11) over these parameters. Unfortunately, the parametrization of orthogonal matrices is difficult and this may not lead to a reliable solution. Instead, the probabilities  $\alpha_{ij}$  will be treated as hidden variables within an EM algorithm. Given the above formulation, EM will estimate both  $Q$  and the  $\alpha_{ij}$ 's, i.e., its output

consists in the alignment of the two embedded spaces as well as a probabilistic assignment between the two sets of points. The latter must be further processed to yield a one-to-one assignment. First we give an outline of algorithm and then we analyse in detail each one of its steps.

Step 0: *Initialization.* Set initial values for the matrix  $Q$  and for the variance  $\sigma$ ;

Step 1: *Expectation.* Compute  $\alpha_{ij}$  and then  $\beta_j$  and  $w_j$  for all  $1 \leq i \leq n$  and  $1 \leq j \leq m$ ;

Step 2: *Maximization.* Use eq. (11) to find a new estimate for  $Q$ . Allow the variance to decrease geometrically with  $\sigma^{\text{new}} = \kappa\sigma$  and  $0 < \kappa < 1$ ;

Step 3: *Test.* If the entries of  $Q$  are stabilized or if  $\sigma^{\text{new}} \leq \sigma_{\min}$  then go to step 4, else go to step 1;

Step 4: *One-to-one assignment.* Extract one-to-one assignments from the probabilities  $\alpha_{ij}$  and from the weights  $\beta_j$ .

**Intialization.** This is a crucial step since the problem that we want to solve has an intrinsic combinatorial nature. The  $k \times k$  covariance matrices  $\Sigma_x$  and  $\Sigma_y$  associated with the embedded point sets  $\mathbf{x}, \mathbf{y}$  characterize the modes of the embedded shape and can be used as a hint to find the most plausible orderings for the eigenvectors. This strategy remains however a heuristic since some shapes have intrinsic identical elongations along several of its modes. We are still left with the sign-reversal ambiguity. Currently we hypothesize over all  $2^k$  for each plausible swap and we evaluate each one using a nearest neighbor test that selects  $\mathbf{x}_i/\mathbf{y}_j$  pairs. Next we estimate the matrix  $Q$  associated with the best pairing.

**Expectation.** This is a classical step that consists in evaluating the probabilities of each possible point pair. This can be relatively slow because it is necessary to compute the Euclidean distance between each  $\mathbf{x}_i$  and each  $Q\mathbf{y}_j$ . We note that this is a classical problem with which K-means and EM algorithms are faced. See [3] for a discussion on efficient implementation of this step.

**Maximization.** The maximization of the expectation, i.e., eq. (8) is replaced in our case by the a minimization, as already explained above. The matrix  $Q$  to be estimated belongs to the group of orthogonal matrices. Unlike its subgroup of rotations (which is a Lie group), the orthogonal group does not yield a parameterization. Nevertheless, eq. (11) does have a closed-form solution. Indeed, we follow the closed-form solution for rotation recovery in [1] and

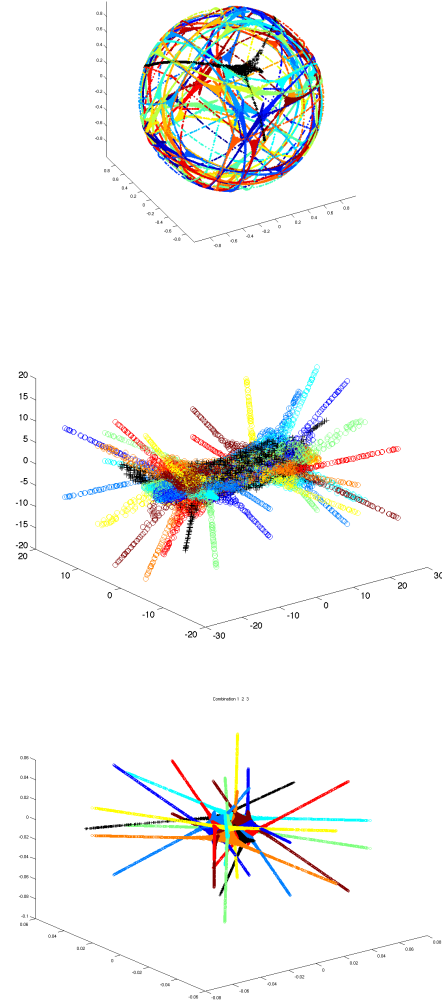


Figure 2. Algorithm initialization. There are 8 possible alignments but none of them is truly satisfactory: Laplacian embedding (top), Isomap (middle), and LLE (bottom).

[25] while relaxing the rotation constraint in order to enable the reflections to be recovered. In brief: First we compute the matrix  $H_{k \times k} = \sum_j \beta_j w_j \mathbf{y}_j \mathbf{y}_j^\top$ . Second we compute the singular value decomposition  $H = U \Lambda V^\top$ . Third we retain the solution  $Q = V U^\top$ .

Formally, one should include the evaluation of  $\sigma$  in the optimization process. It was observed by many authors that when the variance is estimated like this, it decreases too quickly and the algorithm can be trapped in a local minimum. We start with a large variance and we decrease it geometrically, as explained above. The variance should not be smaller than the variance of the expected measurement noise of the observations. This annealing process is crucial

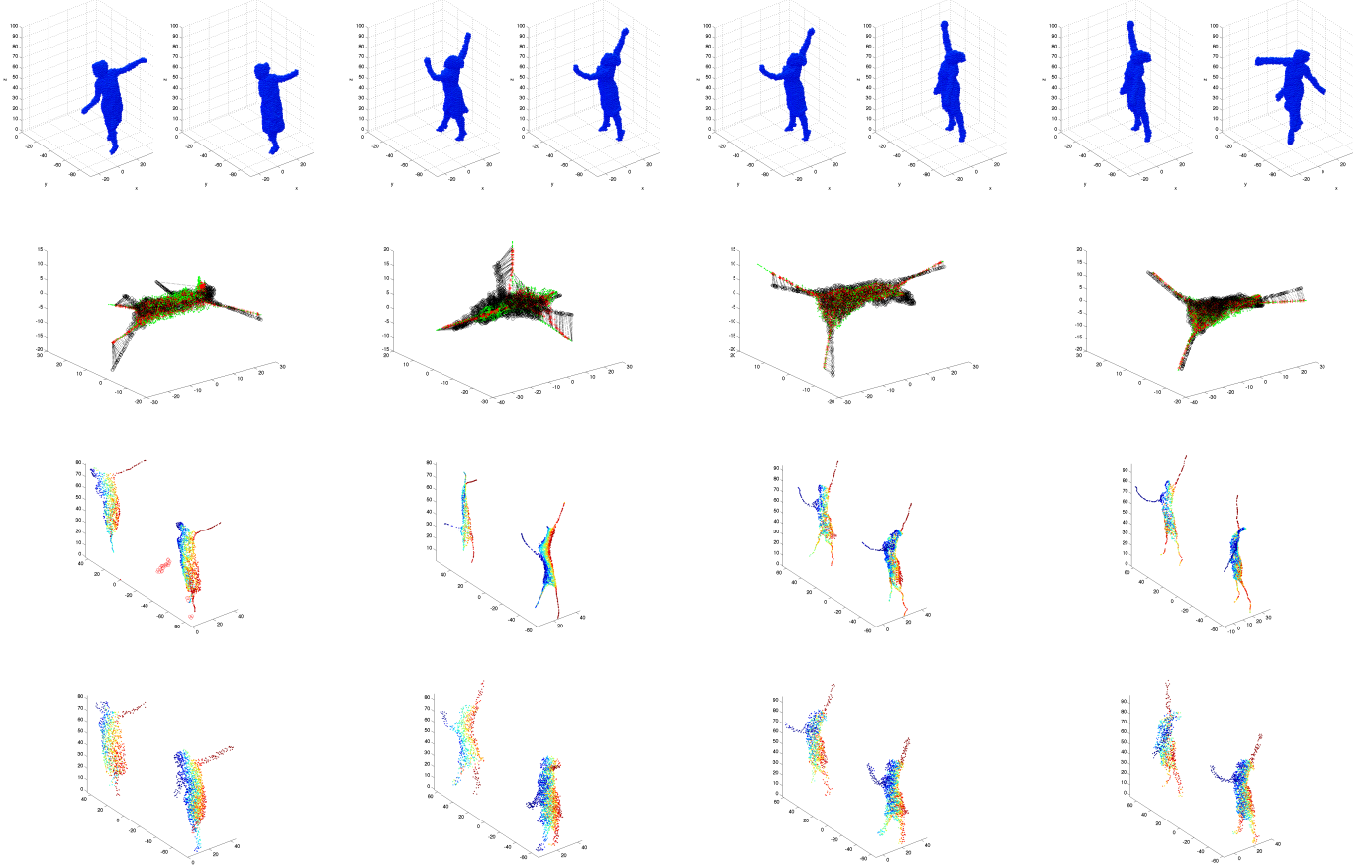


Figure 3. This figure shows the result of matching frames 1/6,21/26,26/31 and 31/36 with Isomap. The first row shows the initial voxel representation. The second row shows the final result of matching in embedded space (here we chose  $k = 3$  for visualization purposes). The third row shows the output of the EM algorithm, and the fourth row shows the result of the final one-to-one assignment between the sets of points. Matching is codified by color. In this case the algorithm successfully matched 90-95% of the points. Note the missing arm in the first column correctly identified as matching outliers (red disconnected circles) in the EM step.

to both escape from local minima and eliminate outliers.

**One-to-one assignment.** This is the final step of the algorithm. The many-to-many probabilistic registration found so far must be interpreted in terms of a one-to-one deterministic association between two largest subsets of the two sets of points. First, for each point  $\mathbf{x}_i$  we select its best assignment, i.e.,  $z_i = \mathbf{y}_j$  is found with  $\arg \max_j \alpha_{ij}$ . Second, we select inliers. The expression of  $\alpha_{ij}$  given by eq. (7) is crucial. Notice that the probability of  $\mathbf{x}_i$  to be an inlier is in the interval  $[0, \frac{1}{1+\vartheta_C}]$ . Remember that the *outlier constant*  $\vartheta_C$  stems from the choice of a uniform probability distribution to describe outliers. In all our experiments we set  $\vartheta_C = 1$ . This prevents our algorithm both from rejecting too many assignments and from running into numerical instabilities when  $\sum_j \exp(-d^2(\mathbf{x}_i, Q\mathbf{y}_j)/2\sigma^2)$  in eq. (7) is close to zero. We adopt the following outlier rejection deci-

sion ( $0 < \nu < 1$ ): If  $\alpha_{ij} > \frac{\nu}{1+\vartheta_C}$  then  $\mathbf{x}_i$  is an inlier, else it is an outlier.

So far we selected a unique assignment for each inlier  $\mathbf{x}_i^{\text{in}}$ , but there may be several inliers associated with the same point  $\mathbf{y}_j$ . In order to obtain a one-to-one correspondence, we take the expectation over all inliers associated with  $\mathbf{y}_j$ :  $w_j = \sum_i \alpha_{ij} \mathbf{x}_i^{\text{in}} / \sum_i \alpha_{ij}$ .

## 5. Experiments

We applied the algorithm to a human-motion sequence. The dancer performs relatively fast motions. We used six cameras and a standard space-carving algorithm in order to acquire a voxel representation at 30 frames per second. The matching results shown below were obtained by applying the algorithm to frames 1-to-6, 21-to-26, 26-to-31, and 31-to-36. Initially there are approximately 5000 voxels at each



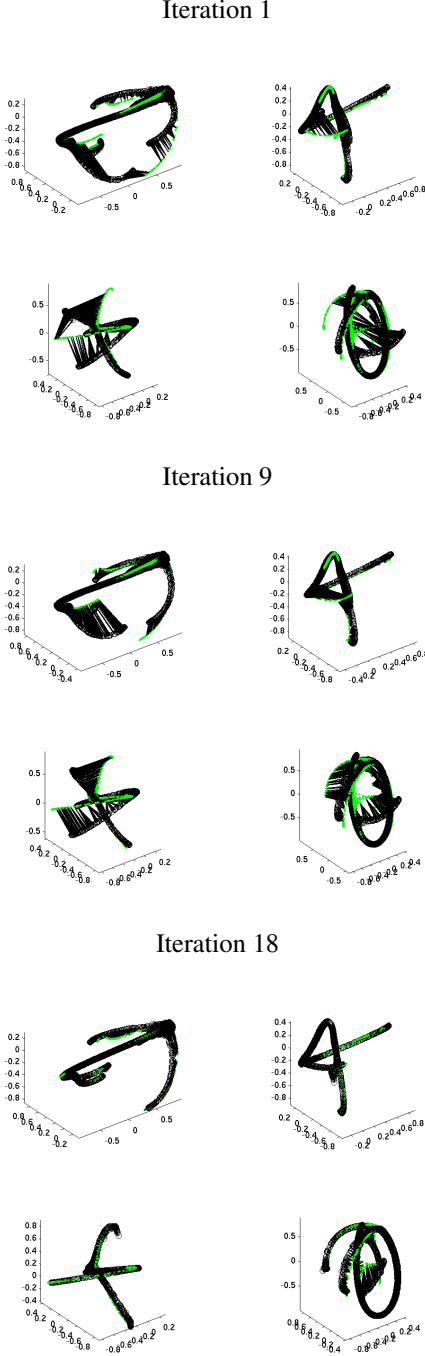


Figure 4. Three steps of the evolution of the EM algorithm in the case of Laplacian embedding and for  $k = 6$ . Matching and alignment of the black and the green embedded shapes. The alignment is projected onto four subspaces of dimension 3 and matchings are indicated by dotted lines.

time step. We reduced these data to obtain 2000-3000 3-D points at each time step.

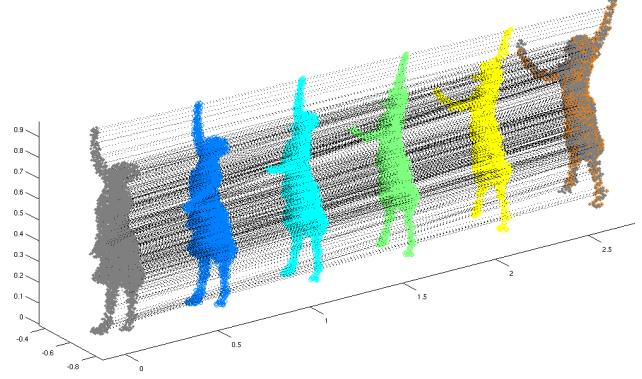


Figure 5. Interpolation. Pointsets of the dancer at different frames are matched and then interpolated creating 4 new intermediate frames. The leftmost and rightmost gray pointsets correspond to the original poses. Colored poses are the result of the interpolation. Dotted lines display some of the associations

We used our method in conjunction with three embedding techniques: Laplacian embedding, Isomap, and LLE. Figure 2 shows the result of the initialization step described above for the three methods and for  $k = 3$ . In this case there are 8 possible alignments.

Figure 3 shows the result of matching using Isomap. Notice the large arm motions of the dancer. The algorithm successfully matched 90-95% of the points.

Figure 4 shows the evolution of the EM algorithm when matching two sets of points from the same sequence using Laplacian embedding and with  $k = 6$ . The figure shows alignment obtained iteratively and projected on four subspaces of dimension 3. In this case the algorithm checked 64 sign ambiguities at initialization and then converged in 18 EM iterations (there are 46080 possible solutions).

Finally Figure 5 illustrates one of the applications of our method: the interpolation of two frames. Here two different poses are matched, and the result is used to interpolated 4 intermediate frames.

## 6. Conclusion

We addressed the problem of matching two articulated shapes. Since we adopted a probabilistic framework, each shape is described by a set of points. Although linear, locally linear, and non-linear methods to build embedded descriptions of such point sets are well developed, the problem of aligning embedded representations has not been properly addressed. We analysed the ambiguities associated with embedded alignment, we revealed and analysed the combinatorial nature of the task. We showed that the problem can be cast into finding a set of point correspondences



that minimize an orthogonal transformation. We described in detail an EM algorithm that starts with many-to-many assignments, converges to many-to-one assignments, and eventually leads to one-to-one assignments. The method was successfully applied to large sets of points corresponding to a human-motion sequence. In the future we plan to analyse more thoroughly the initialization step of our algorithm and to apply it to more general graph-isomorphism problems.

## References

- [1] K. S. Arun, T. S. Huang, and S. D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-9(5):698–700, September 1987.
- [2] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering, 2002.
- [3] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [4] T. Caelli and S. Kosinov. An eigenspace projection clustering method for inexact graph matching. *Pattern Analysis and Machine Intelligence PAMI*, 2004.
- [5] M. Carcassoni and E. R. Hancock. Correspondence matching with modal clusters. *Pattern Analysis and Machine Intelligence PAMI*, 25(12):1609–1615, 2003.
- [6] M. Carcassoni and E. R. Hancock. Spectral correspondence for point pattern matching. *Pattern Recognition*, 36:193–204, 2003.
- [7] F. Chang. *Spectral Graph Theory*. American Mathematical Society, 1997.
- [8] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. 89(2-3):114–141, February 2003.
- [9] A. Cross and E. Hancock. Graph matching with a dual-step EM algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1236–1253, November 1998.
- [10] G. Dewaele, F. Devernay, R. Horaud, and F. Forbes. The alignment between 3-d data and articulated shapes with bending surfaces. In *Proceedings of the 9th European Conference on Computer Vision, Graz, Austria*, volume III of LNCS, pages 578–591. Springer, May 2006.
- [11] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *Pattern Analysis and Machine Intelligence PAMI*, 25(10):1285–1295, 2003.
- [12] S. Gold and A. Rangarajan. A Graduated Assignment Algorithm for Graph Matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):377–388, 1996.
- [13] S. Granger and X. Pennec. Multi-scale em-icp: A fast and robust approach for surface registration. In *ECCV02*, page IV: 418 ff., 2002.
- [14] R. Horaud and H. Sossa. Polyhedral object recognition by indexing. *Pattern Recognition*, 28(12):1855–1870, 1995.
- [15] V. Jain and H. Zhang. Robust 3-d shape correspondence in the spectral domain. In *IEEE International Conference on Shape Modeling and Applications 2006*, 2006.
- [16] B. Luo and E. Hancock. A unified framework for alignment and correspondence. *CVIU*, 92(1):26–55, October 2003.
- [17] J. Mejia, T. Villela, and J. Braga. The CCD stellar sensor of the masco telescope pointing system. *Advances in Space Research*, 26(9):1407–1410, 2000.
- [18] M. Niethammer, M. Reuter, F. Wolter, N. P. S. Bouix, M. Koo, and M. Shenton. Global medical shape analysis using the laplace-beltrami spectrum. In *MICCAI07, 10th International Conference on Medical Image Computing and Computer Assisted Intervention*, 2007.
- [19] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.
- [20] G. Scott and C. L. Higgins. An algorithm for associating the features of two images. In *Biological Sciences*, volume 244, pages 21–26, 1991.
- [21] L. Shapiro and J.M.Brady. Feature-based correspondence an eigenvector approach. In *Image and Vision Computing*, volume 10, pages 283–288, 1992.
- [22] W. Taylor. Protein structure comparison using bipartite graph matching and its applications to protein structure classification. *Molecular and Cellular Proteomics*, pages 334–340, 2002.
- [23] J. Tenenbaum, V. de Silva, and J. Jangford. A global geometric framework for non-linear dimensionality reduction. *Science*, 290:2319–2323, 2000.
- [24] S. Umeyama. An eigendecomposition approach to weighted graph matching problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):695–703, May 1988.
- [25] S. Umeyama. Least-squares estimation of transformation parameters between two point pattern. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, April 1991.
- [26] W. Wells III. Statistical approaches to feature-based object recognition. *International Journal of Computer Vision*, 28(1/2):63–98, 1997.
- [27] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal on Computer Vision*, 13:119–152, 1994.